GEOG0113: GEOGRAPHY IN THE FIELD

INTRODUCTION TO STATISTICS

GETTING YOU STARTED WITH THE LEARNING MATERIALS

Dr Anwar Musah (a.musah@ucl.ac.uk)
Lecturer in Social and Geographic Data Science
UCL Geography

How access the learning materials?

Week 4: Introduction to statistics

Overview

Welcome to **Introduction to Statistics**. So far you have been introduction to the core tenets and principles of geography taught through the different themes touching on **physical sciences** (e.g., coastal and rivers) and **social sciences** (e.g., ethnographic methods and urbanism), and now, in Week 4's session, it will focuses on the themes of **statistics** (i.e., data analysis). In week 4, we provide you with an introduction to basic statistical methods for exploring secondary data from Barcelona, using a programming software tool known as **R/RStudio**. These self-guided tutorials will be prescriptive with instructions supported by with guidance videos .

These tutorials will prepare you for the group workshop which will facilitated by the PGTAs - the workshop will the opportunity for you to ask questions and support so as you can work on you **Data Analysis Worksheet** independently which is worth 10% of the assessment.

Materials and activities

All information concerning room locations, as well as learning resources for Week 4 are hosted on a dedicated webpage: please click on **[GEOG0013: Introduction to Statistics]** to access it**. Please read through the "Welcome" chapter and work through the instructions in sections: 1.) Getting Started with RStudio; 2.) Basics of Managing Data in RStudio; and 3.) Exploratory Analysis in RStudio. Good luck!**

Further resources

Forums: Questions and Discussions

Mark as done

Use this platform to post general questions about the content on this webpage. You are welcome to post your problems on technical issues encountered. The PGTAs will respond accordingly with solutions. Students are welcome to engage and support other students with solutions through this forum during this self-study period.

Welcome the GEOG0013: Introduction to Statistics webpage

GEOG0013: Geography in the Field 1 (GIF1) 2022/23

Welcome to Introduction to Statistics

Welcome to **GIF1's Introduction to Statistics**. This module has been designed as an introduction to the core tenets and principles of geography taught through many different themes touching on **physical sciences** (e.g., coastal and rivers), **social sciences** (e.g., ethnographic methods and urbanism) and **statistics** (e.g., data analysis and GIS). This session focuses on the themes of **statistics** (i.e., data analysis). In week 4, we provide you with an introduction to basic statistical methods for exploring & representing geographic secondary data from Barcelona, using a programming software tool known as **R/RStudio**.

Structure

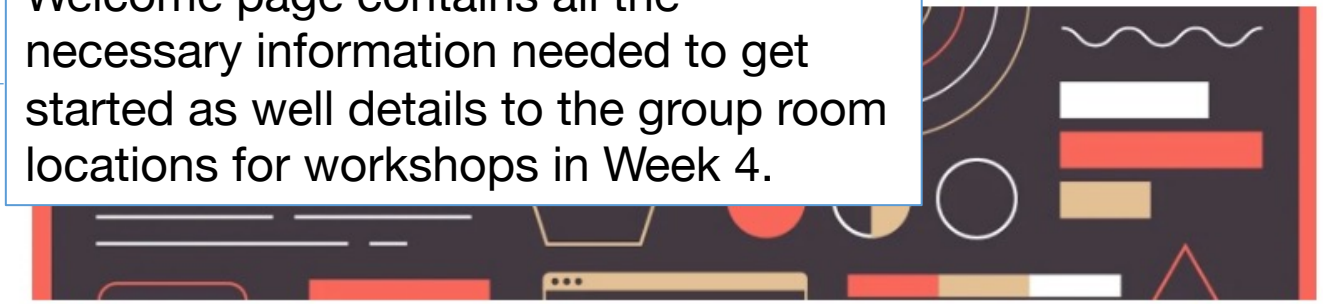
The structure of this GIF1 component is a 3-hour group computer practical workshop that will be delivered in-person. All students have been allocated to one of 8 groups (i.e., group **A**, **B**, **C**, **D**, **E**, **F**, **G**, **H** or **I**). The workshop for week 4's **Introduction to Statistics** will take place on the following dates:

1. **Wednesday 26th October, 09:00am-12:00pm** for groups **F**, **G**, **H**, **I**
2. **Friday 28th October, 02:00pm-05:00pm** for groups **A**, **B**, **C**, **D**, **E**

Navigating through learning materials? [1]

Welcome page contains all the necessary information needed to get started as well details to the group room locations for workshops in Week 4.

GEOG0013: Geography in the Field 1 (GIF1) 2022/23



Welcome to **GIF1's Introduction to Statistics**. This module has been designed as an introduction to the core tenets and principles of geography taught through many different themes touching on **physical sciences** (e.g., coastal and rivers), **social sciences** (e.g., ethnographic methods and urbanism) and **statistics** (e.g., data analysis and GIS). This session focuses on the themes of **statistics** (i.e., data analysis). In week 4, we provide you with an introduction to basic statistical methods for exploring & representing geographic secondary data from Barcelona, using a programming software tool known as **R/RStudio**.

Structure

The structure of this GIF1 component is a 3-hour group computer practical workshop that will be delivered in-person. All students have been allocated to one of 8 groups (i.e., group **A**, **B**, **C**, **D**, **E**, **F**, **G**, **H** or **I**). The workshop for week 4's **Introduction to Statistics** will take place on the following dates:

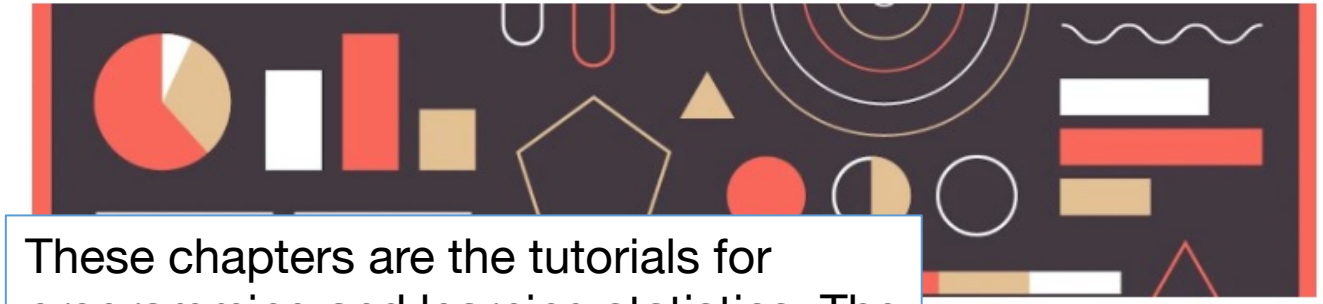
1. **Wednesday 26th October, 09:00am-12:00pm** for groups **F**, **G**, **H**, **I**
2. **Friday 28th October, 02:00pm-05:00pm** for groups **A**, **B**, **C**, **D**, **E**

Navigating through learning materials? [2]

These chapters are the tutorials for programming and learning statistics. The contain the instructions and support videos.

GEOG0013: Geography in the Field 1 (GIF1) 2022/23

Welcome to Introduction to Statistics



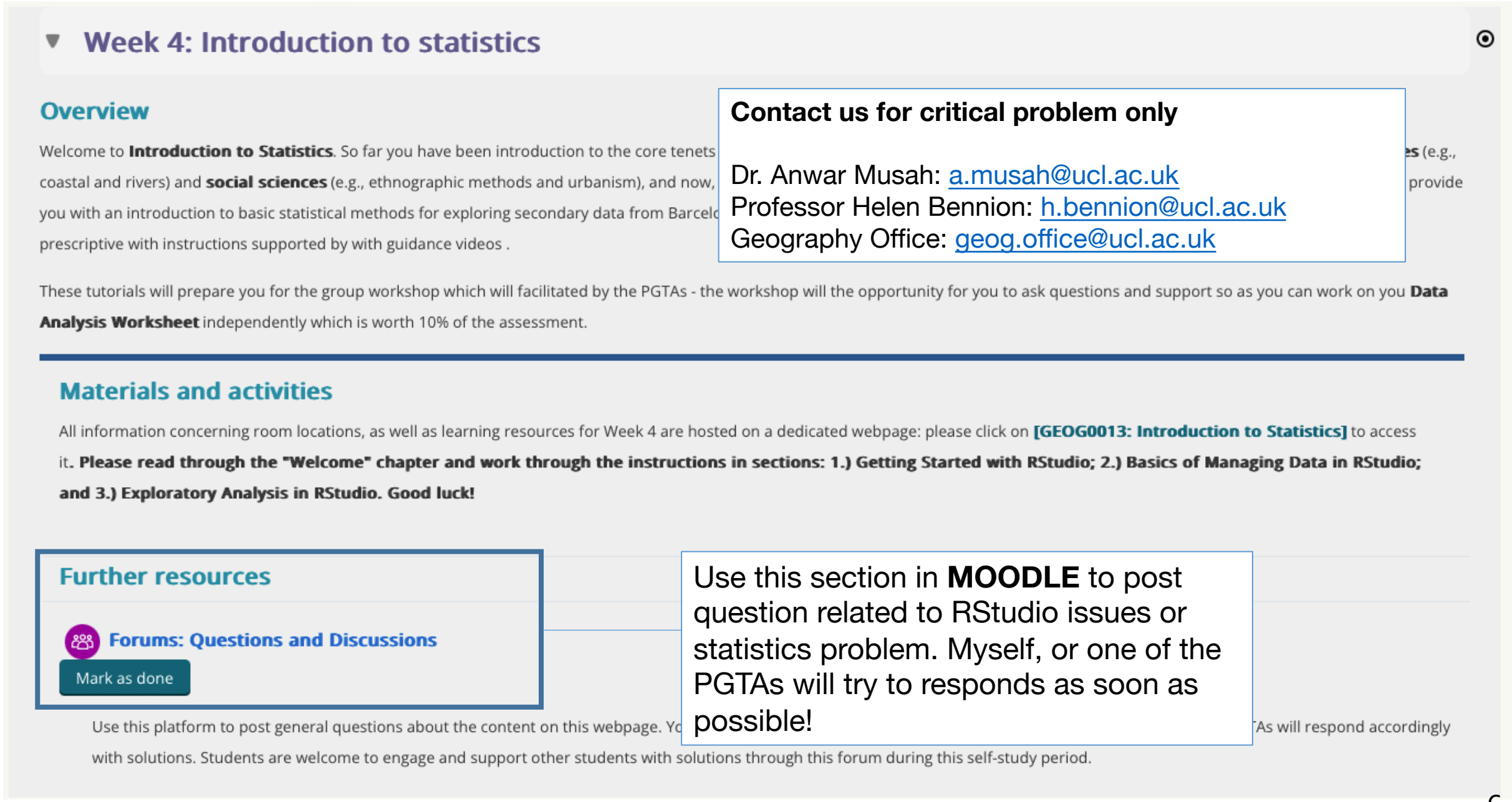
signed as an introduction to the
t themes touching on **physical**
ic methods and urbanism) and **statistics** (e.g., data analysis and GIS). This session focuses on the themes of **statistics** (i.e., data analysis). In week 4, we provide you with an introduction to basic statistical methods for exploring & representing geographic secondary data from Barcelona, using a programming software tool known as **R/RStudio**.

Structure

The structure of this GIF1 component is a 3-hour group computer practical workshop that will be delivered in-person. All students have been allocated to one of 8 groups (i.e., group **A**, **B**, **C**, **D**, **E**, **F**, **G**, **H** or **I**). The workshop for week 4's **Introduction to Statistics** will take place on the following dates:

1. **Wednesday 26th October, 09:00am-12:00pm** for groups **F**, **G**, **H**, **I**
2. **Friday 28th October, 02:00pm-05:00pm** for groups **A**, **B**, **C**, **D**, **E**

How to post questions and key contacts



Week 4: Introduction to statistics

Overview

Welcome to **Introduction to Statistics**. So far you have been introduction to the core tenets ...es (e.g., coastal and rivers) and **social sciences** (e.g., ethnographic methods and urbanism), and now, ... provide you with an introduction to basic statistical methods for exploring secondary data from Barcelo... prescriptive with instructions supported by with guidance videos .

These tutorials will prepare you for the group workshop which will facilitated by the PGTAs - the workshop will the opportunity for you to ask questions and support so as you can work on you **Data Analysis Worksheet** independently which is worth 10% of the assessment.

Materials and activities

All information concerning room locations, as well as learning resources for Week 4 are hosted on a dedicated webpage: please click on **[GEOG0013: Introduction to Statistics]** to access it**. Please read through the "Welcome" chapter and work through the instructions in sections: 1.) Getting Started with RStudio; 2.) Basics of Managing Data in RStudio; and 3.) Exploratory Analysis in RStudio. Good luck!**

Further resources

Forums: Questions and Discussions

Mark as done

Use this platform to post general questions about the content on this webpage. Yo... TAs will respond accordingly with solutions. Students are welcome to engage and support other students with solutions through this forum during this self-study period.

Contact us for critical problem only

Dr. Anwar Musah: a.musah@ucl.ac.uk
Professor Helen Bennion: h.bennion@ucl.ac.uk
Geography Office: geog.office@ucl.ac.uk

Use this section in **MOODLE** to post question related to RStudio issues or statistics problem. Myself, or one of the PGTAs will try to responds as soon as possible!

How to post questions and key contacts

Week 4: Introduction to statistics

Overview

Welcome to **Introduction to Statistics**. So far you have been introduction to the core tene…ces (e.g., coastal and rivers) and **social sciences** (e.g., ethnographic methods and urbanism), and nov…ve provide you with an introduction to basic statistical methods for exploring secondary data from Barcelona, using a programming software tool known as **R/RStudio**. These self-guided tutorials will be prescriptive with instructions supported by with guidance videos .

These tutorials will prepare you for the group workshop which will facilitated by the PGTAs - the workshop will the opportunity for you to ask questions and support so as you can work on you **Data Analysis Worksheet** independently which is worth 10% of the assessment.

Contact us for critical problem only

Dr. Anwar Musah: a.musah@ucl.ac.uk
Professor Helen Bennion: h.bennion@ucl.ac.uk
Geography Office: geog.office@ucl.ac.uk

Materials and activities

All information concerning room locations, as well as learning resources for Week 4 are hosted on a dedicated webpage: please click on **[GEOG0013: Introduction to Statistics]** to access it**. Please read through the "Welcome" chapter and work through the instructions in sections: 1.) Getting Started with RStudio; 2.) Basics of Managing Data in RStudio; and 3.) Exploratory Analysis in RStudio. Good luck!**

Further resources

Forums: Questions and Discussions

Mark as done

Use this platform to post general questions about the content on this webpage. Yo…TAs will respond accordingly with solutions. Students are welcome to engage and support other students with solutions through this forum during this self-study period.

Use this section in **MOODLE** to post question related to RStudio issues or statistics problem. Myself, or one of the PGTAs will try to responds as soon as possible!

Format and learning flow [1]

To generate `bmi` into our data frame, we would need to access the `height` (m) and `weight` (kg) columns using the `$` from the data frame its stored to, and apply the above formula as a code to generate the new `bmi` column:

```
# Create 'bmi' in the data frame i.e.,'dataset' and calculate 'bmi'
# using the $weight and $height
dataset$bmi <- dataset$weight/((dataset$height)^2)
# View the data frame 'dataset' and you will see the new bmi variable inside
View(dataset)
```

You can overwrite the `height` (m) column to change its units into centimeters by multiplying it to 100; equally, the `weight` (kg) column can be overwritten and converted from units of kilograms to grams by multiplying it to 1000.

```
# using $height and *100
dataset$height <- dataset$height*100
# using $weight and *100
dataset$weight <- dataset$weight*1000
# use View() the data frame 'dataset' and you will see the updated variables
View(dataset)
```

There will always be a video to explain the necessary theory of the statistical method, as well as an explanation of the code

There will always be a **text** to explain the necessary what the code is also doing, and steps for coding – so here, you will have to read through the **text.**

There will always be **code chunks** for you to replicate and study in your own RStudio.

Format and learning flow [2]

To generate `bmi` into our data frame, we would need to access the `height` (m) and `weight` (kg) columns using the `$` from the data frame its stored to, and apply the above formula as a code to generate the new `bmi` column:

```
# Create 'bmi' in the data frame i.e.,'dataset' and calculate 'bmi'
# using the $weight and $height
dataset$bmi <- dataset$weight/((dataset$height)^2)
# View the data frame 'dataset' and you will see the new bmi variable inside
View(dataset)
```

You can overwrite the `height` (m) column to change its units into centimeters by multiplying it to 100; equally, the `weight` (kg) column can be overwritten and converted from units of kilograms to grams by multiplying it to 1000.

```
# using $height and *100
dataset$height <- dataset$height*100
# using $weight and *100
dataset$weight <- dataset$weight*1000
# use View() the data frame 'dataset' and you will see the updated variables
View(dataset)
```

There will always be a **video** to explain the necessary theory of the statistical method, as well as an explanation of the code

There will always be a text to explain the necessary what the code is also doing, and steps for coding – so here, you will have to read through the text.

There will always be code chunks for you to replicate and study in your own RStudio.

Format and learning flow [3]

To generate `bmi` into our data frame, we would need to access the `height` (m) and `weight` (kg) columns using the `$` from the data frame its stored to, and apply the above formula as a code to generate the new `bmi` column:

```
# Create 'bmi' in the data frame i.e.,'dataset' and calculate 'bmi'
# using the $weight and $height
dataset$bmi <- dataset$weight/((dataset$height)^2)
# View the data frame 'dataset' and you will see the new bmi variable inside
View(dataset)
```

You can overwrite the `height` (m) column to change its units into centimeters by multiplying it to 100; equally, the `weight` (kg) column can be overwritten and converted from units of kilograms to grams by multiplying it to 1000.

```
# using $height and *100
dataset$height <- dataset$height*100
# using $weight and *100
dataset$weight <- dataset$weight*1000
# use View() the data frame 'dataset' and you will see the updated variables
View(dataset)
```

There will always be a **video** to explain the necessary theory of the statistical method, as well as an explanation of the code

There will always be a **text** to explain the necessary what the code is also doing, and steps for coding – so here, you will have to read through the **text.**

There will always be code chunks for you to replicate and study in your own RStudio.

The code chunks will be annotated with a # (hash tag) with a comment to give context. So this is not actual code ■

The real code is the code chuck with no # (hash tag) at the beginning. ■

Week 4's Practical Workshop session

1. **Wednesday 26th October, 09:00am-12:00pm** for groups **F**, **G**, **H**, **I**
2. **Friday 28th October, 02:00pm-05:00pm** for groups **A**, **B**, **C**, **D**, **E**

Group room allocations: Here are the details for the room location. Please use the table below view the map details of the location, as well as watch the videos embedded them to see how to get to the location. To avoid confusion, please go to your group's designated cluster room.

Group	Date	Location
A	**Friday, 28/10**	Chadwick Building, Room 2.23, Map; **Video**
B	**Friday, 28/10**	Bedford Way (26), Room G11, Map; **Video**
C	**Friday, 28/10**	ICH Wolfson Centre Room I (1st floor), **Map**; **Video**
D	**Friday, 28/10**	Birkbeck Gordon Sq (43); Room B06, **Map**; **Video**
E	**Friday, 28/10**	IOE Bedford Way (20), Room 784, Map; **Video**
F	**Wednesday, 26/10**	Birkbeck Gordon Sq (43), Room B06, **Map**; **Video**
G	**Wednesday, 26/10**	North-West Wing, Room G17, Map; **Video**
H	**Wednesday, 26/10**	Birkbeck Gordon Sq (43), Room 122, **Map**; **Video**
I	**Wednesday, 26/10**	IOE Bedford Way (20), Room 784, Map; **Video**

IMPORTANT NOTE: All students on the Wednesday workshops, please bring your own laptops with you to the computer practicals. If you do not own a laptop, or not in possession of one, you can use **UCL Laptop Loan services [CLICK HERE]**.

Try to complete as much of the self-guide tutorials on your own before heading into the Workshop.

The goal is to learn RStudio and the statistics to prepare for the assignment.

The Data Analysis assignment sheets will be released on Tuesday.

You can use the workshop time to ask PGTAs to rectify any problems you encountered. They may, or may stand in front of blackboard and teach. It's your responsibility to come in prepared.

Those in the following groups: F, G, H and I

Must bring your own laptops for Wednesday.

Any questions?

a.musah@ucl.ac.uk